Journal of Geodetic Science

• 2(3) • 2012 • 188-199 DOI: 10.2478/v10156-011-0039-2 •

Modeling regional sea level rise using local tide gauge data

Research Article

H. Bâki Iz[1]*, L. Berry[2] and M. Koch[2]

1 *Dept. of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China*

2 *Center for Environmental Studies and Biological Sciences Dept., Florida Atlantic University, Boca Raton, FL, USA*

Abstract:

Currently regional mean sea level trends and variations are inferred from the analysis of several individual local tide gauge data that span only a long period of time at a given region. In this study, we propose using a model to merge various tide gauge data, regardless of their time span, in a single solution, to estimate parameters representative of regional mean sea level trends. The proposed model can account for the geographical correlations among the local tide gauge stations as well as serial correlations, if needed, for individual stations' data. Such a vigorous regional solution enables statistically optimal uncertainties for estimated and projected trends. The proposed formulation also unifies all the local reference levels by modeling their offsets from a predefined station's reference level. To test its effectiveness, the proposed model was used to investigate the regional mean sea level variations for the coastal areas of the Florida Panhandle using 26 local tide gauge stations that span approximately 830 years of monthly averages from the Permanent Service for Mean Sea Level repository. The new estimate for the regional trend is 2.14 mm/yr with a ±0.03 mm/yr standard error, which is an order of magnitude improvement over the most recent mean sea level trend estimates and projections for the Florida region obtained from simple averages of local solutions.

Keywords:

Florida sea level change • Mean sea level regional sea level • Tide gauge model

Received 10-06-2012; accepted 20-08-2012

1. Introduction

Long term changes in the mean sea level (MSL) impacts shoreline and beach erosion, coastal and wetlands inundation, storm surge flooding and coastal development. Trends in sea level have traditionally been calculated from long term data sets at a few locations. However, inferences from local tide gauge series are not representative of regional mean sea level variations because there are only few tide gauge stations with long records (longer than 60 years, Douglas 1991[1]) around the world, and their uncertainties prohibit accurate regional sea level projections (Mitchum 2011).

Msl trend estimates from different local tide gauge measurements with mixed record lengths are rarely averaged to estimate regional trend in the MSLs because the local mean sea level estimates from shorter records are corrupted by interdecadal fluctuations (Gornitz et al. 1982 and Barnett 1984) and by unmodeled local effects (Iz and Ng 2005). This practice leaves out considerable amount of tide gauge data from the analysis, creating an under sampling problem similar to in estimating global averages (Cabanes et al. 2001), which may otherwise contribute significantly in estimating sea level trends and variations.

Moreover, nearby tide gauges with similar record lengths may exhibit high geographical correlations hence they are used only for quality control (the buddy analysis by the Permanent Mean Sea Level Service, PSMSL, 2012). The nearby stations are often discarded because it is suggested that these stations' data do not contain any new information, yet their inclusion may still improve the solution statistics as will be demonstrated in our study.

Currently regional models for estimating MSL variations using tide

*E-mail: H.Baki.Iz@gmail.com

[1] *This estimate is optimistic. In some cases, unmodeled local and regional effects cannot be eliminated or reduced effectively using TG data that span over 60 years (Iz 2006).*

VERSITA

gauge data involve simple or weighted averaging local solutions' trends, (Douglas 1991, Mazzotti et al. 2008, respectively). Regional solutions are avoided because the estimates may be biased if the geographical (spatial) correlations among the nearby stations are omitted and cause over/under estimation of the uncertainties of the estimated regional parameters.

The disturbances of the tide gauge data may also exhibit temporal correlations (autocorrelations, the red noise for the first order autoregressive process) that need to be accounted for in local as well as in potential regional solutions. Their omission may cause underestimation of the error estimates of the solution parameters, thereby leading potential Type I errors in null-hypothesis testing for the significance of the model parameters. Currently the effect of the serial correlation of the residuals is accounted for by using an inflation factor in satellite altimetry derived trend estimates (Church et al. 2011). No studies have been carried out to assess their impact on trend estimates from the local tide gauge data.

Hence, this study addresses the following issues:

1. How to combine/model various local tide gauge data to estimate regional MSL rise and variations for short and long records irrespective of different local tide gauge reference levels?

2. How to estimate and subsequently model spatially correlated station disturbances among the local stations' tide gauge data in a regional solution?

3. How to account for serial autocorrelations of each local tide gauge time series records in a regional solution for more realistic error statistics for the estimated trend parameters? and,

4. Quantify, assess and validate the impact of a vigorous regional model in estimating sea level changes using multiple station data as opposed to the current practice of simple pooling of local solutions for the tide gauge stations around the Florida Panhandle.

The following sections will first investigate local solutions for each station in a region to estimate model parameters using ordinary least squares method. These results will be used as baselines for subsequent solutions. The solution residuals will be scrutinized for a follow up regional model solution. The residuals will also be used to estimate geographical correlations among the participating stations as well as the serial correlation for each station's disturbances. In the second step, geographical and serial correlation information will be used to establish a full variance-covariance matrix of the errors, which will be deployed in a generalized least squares solution for estimating the regional means sea level trend parameters.

A trend estimate in a regional solution is statistically optimal, thanks to the rigorous formulation, as should be evidenced by the improved error statistics for the trend estimate, its standard error and improved R^2 (*adj*) statistic (Neter et al. 1996) as opposed to

Figure 1. Distribution of 26 tide gauge stations in Florida State (PSMSL data catalogue on Google Earth Map).

an estimate that can be obtained by averaging the local trend estimates. The use of aggregated tide gauge data in such combination solution provides additional information that is present in some series but missing in others because of their localities, thus better in capturing regional MSL variations.

The new model will be tested using 26 local tide gauge stations that span over 830 years (monthly averages) around the Florida Panhandle for the local and regional solutions.

2. Tide Gauge Stations and Data around Florida Panhandle

Permanent Service for Mean Sea Level (PSMSL) repository maintains a Tide gauge database from over 1800 stations since 1933. PSMSL repository offers *Metric and Revised Local Reference* data (PSMSL, 2011). *Metric* data is the raw data directly received from the authorities. The revised local reference data contains monthly and annual MSL data referenced to a common reference level. The reference level is defined 7 m below the global MSL to avoid negative monthly and annual MSL values. Only two thirds of the stations in the PSMSL database, however, have been adjusted to a common reference level. The recent MSL trend analyses use the revised local reference data sets extracted from the PSMSL repository that passed a consistency check.

Figure 1 displays the locations of all the tide gauge stations in the PSMSL repository for the State of Florida, USA. The revised local reference tide gauge data, downloaded in April 2011, span over 10,000 monthly averages (∼830 yr), with Fernandina station data being the longest in the series (112 yr), and Lake Worth, Palm Beach data (3.8 yr) being the shortest series. Despite the preliminary adjustment by the PSMSL, the revised local reference data still exhibit small reference level differences, as it will be shown in the following sections, with Virginia Key's tide gauge station's data having the largest offset as shown in the inset of Fig. 5.

2.1. Local Tide Gauge Model and Local Solutions

The following well-known empirical trigonometric model is used to estimate local MSL trend parameters:

$$y_t = a + b(t-\bar{t}) + \sum_{h=1}^{3}\left[\alpha_h \cos\left(\frac{2\pi}{P_h}\right)(t-\bar{t}) + \right.$$
$$\left. + \gamma_h \sin\left(\frac{2\pi}{P_h}\right)(t-\bar{t})\right] + e_t. \tag{1}$$

In this expression, y_t represents the monthly averaged tide gauge data at an epoch t at a given station. The unknown intercept, a, relates local tide gauge data to the vertical reference level, b represents the local MSL trend to be estimated and $\bar{t}$ refers to the middle epoch of the series. The unknown coefficients of the sine and cosine terms are denoted by α an γ from which the amplitudes and the phase angles of the semi-annual, annual, and nodal (18.6 yr) periods denoted by P_h, are determined.
The *random* variable e_t represents the disturbances in the tide gauge data, the lump-sum effect of random instrument errors and unmodeled effects assumed to be less influential in long tide gauge series data, with the following properties:

$$\mathrm{E}[e_t] = 0, \quad \mathrm{Var}[e_t] = \sigma^2. \tag{2}$$

The above relationships also assume that the disturbances are serially independent as common practice in modeling local tide gauge data (this assumption will be assessed in the subsequent section) and σ^2 is the unknown variance of the averaged tide gauge data. The unknown parameters of the formulation are estimated from the local tide gauge data using the ordinary least squares method. Table 1 shows the 17 local trend estimates from 26 tide gauge stations around the Florida Panhandle (Fig. 1) using this model. None of the trend values are corrected for the effect of post glacial rebound, hence the trend estimates refer to the *relative sea-level rise* at each station. The stations experience a negative uplift significantly large in magnitude compared to the sea level trend in this region. Nonetheless the relative uplift rates calculated from the ICE – 5G by Peltier (2004) and Wang ICE5G (2006) models, between stations are about 0.07 mm on the average (Tab. 2). Considering the diminishing differences in the relative rates implied by the newer models, we cannot rule out the null-hypothesis that stations do not experience statistically significant uplift rates with respect to each other in this region. This assumption will also be verified indirectly by the improved regional solutions in this study.
A number of tide gauge data from very close stations (Tab. 1, Fig. 1) are used together in the local solution for better local representations (the same model that will be proposed for the regional scale solution already used here for the combination of very short and very close stations tide gauge data in local solutions).
Table 1 and Fig. 2 show that the magnitude of the estimates are, overall, strongly influenced by the length of the series as expected.

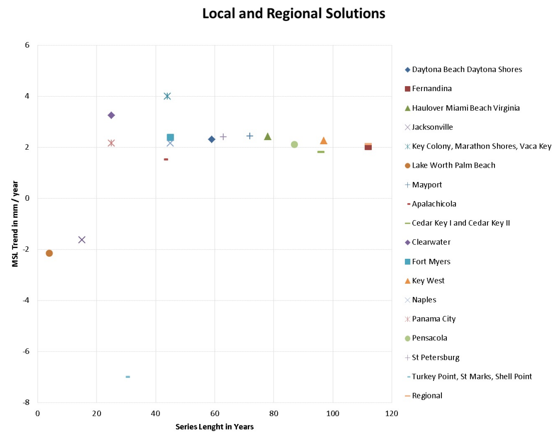

Figure 2. Mean sea level trend estimates for individual local time series and regional solution. Observe the reduced variability with increased series length.

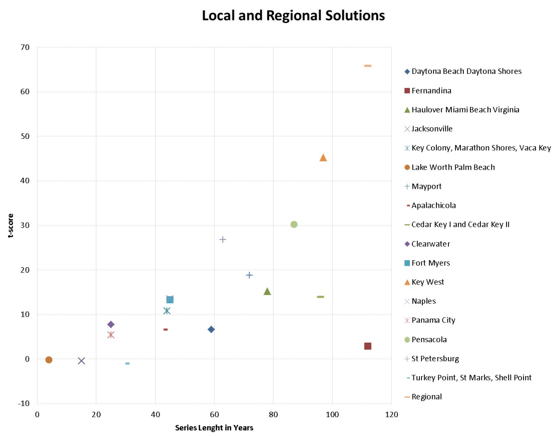

Figure 3. The marked impact of rigorous regional solution is revealed by t-values (signal to noise ratios) calculated by dividing the estimated trends by their standard deviations.

The t-values (trend estimates divided by their standard errors) depicted in Fig. 3 reveal that the corresponding standard errors of the parameters are also correlated with the series lengths; i.e. the longer the series, the smaller the corresponding errors in the trend estimates. This is partly because of the smaller root mean square error of the solutions calculated using larger number of data and partly because of the better separation of the estimates within each local solution due to the length of the series.

The local trend estimates from longer series show less variability and agree well with each other, justifying the wisdom of using longer series in the solutions as proven by Iz (2006). On the other hand, the $R^2(adj)$ values (50 – 79 percent) are not particularly im-

Table 1. Solution statistics for the estimates of the MSL trends from the 26 stations tide gauge data. A number of tide gauge data from very close stations as indicated in the table's solution entries are combined in the local solution in a single series. Local solutions for these data sets take into account the reference level differences. SE stands for Standard Error of the estimate. RMS fit is the root mean square of the residuals.

Solution	Lengthyr	Trendmm/yr	SEmm/yr	RMS Fit(mm)	R^2(adj)
Daytona Beach, Daytona Shores	59	2.31	0.35	82.0	62.6
Fernandina	112	2.03	0.70	80.4	66.4
Haulover, Miami Beach, Virginia	78	2.43	0.16	55.5	98.5
Jacksonville	15	-1.63	3.96	84.4	58.5
Key Colony, Marathon Shores, Vaca Key	44	4.00	0.37	46.7	77.3
Lake Worth, Palm Beach	4	-2.15	13.00	55.2	50.3
Mayport	72	2.44	0.13	76.1	64.7
Apalachicola	43	1.53	0.23	60.7	54.3
Cedar Key I and Cedar Key II	96	1.82	0.13	47.3	75.3
Clearwater	25	3.26	0.42	46.9	73.6
Fort Myers	45	2.40	0.18	52.7	68.1
Key West	97	2.26	0.05	46.4	79.1
Naples	45	2.16	0.16	46.3	70.5
Panama City	25	2.17	0.40	48.9	74.4
Pensacola	87	2.12	0.07	58.0	70.3
St Petersburg	63	2.41	0.09	45.2	76.1
Turkey Point, St Marks, Shell Point	30	-7.00	6.87	58.3	67.0

Table 2. Station uplift rates around Florida Panhandle due to the post glacial rebound (mm/yr) from various models (Peltier 2004, Wang, 2006).

Station	Peltier ICE5G VM2	Peltier ICE5G VM4	Wang ICE5G
Pensacola	-1.93	-1.26	-1.36
Fernandina	-2.13	-1.35	-1.27
Key West	-1.62	-1.27	-1.25

pressive for the local solutions[2]. Although the increasing magnitude of the t-values displayed in Fig. 2, indicative of the positive impact of the data span of the series in the estimation, they do not follow a strict rule. Fernandina station, for instance, with the longest tide gauge data (112 years), has a considerably low t-value mainly caused by the larger residuals which are not explained by the current model parameters (the RMS fit of the Fernandina station is twice as large compared to the RMS fit of the Key West residuals) due to the noisy tide gauge records at this station.

At this point, one might be tempted to estimate the regional trend by averaging the trend estimates of the local stations; which is 1.33 mm/yr with $\pm$0.64 mm/yr for the standard error of the mean. Nonetheless, the impact of the extreme values are evident in the result even with a trimmed mean of 1.72 mm/yr (although the median value of 2.17 mm/yr is a reasonable estimate for the regional average). The corresponding standard errors for the regional estimates by averaging are prohibitively high. The weighted mean of the local rates with weights proportional to the inverse of the variances of the local estimates is 2.61 mm/yr, with a $\pm$0.12 mm/yr standard error of the weighted mean, which may further be refined by removing the extremes. Yet again, this is still not a viable approach because pooling the estimates may not give an optimal solution if the disturbances of the tide gauge measurements are geographically correlated. Moreover, the standard errors of the estimates could be underestimated if the disturbances are serially correlated, which are not accounted for in the ordinary least squares solution. The trend parameters may still be correlated due to the fact that annual semi-annual and node tides do not vary markedly in the region from station to station. The amplitudes of these variations may be different, not because these variations are different at different stations as suggested by Tsimplis and Woodworth (1994), but simply because of the length of the series and other unmodeled local effects unaccounted for in the model, which adversely bias the estimated phase and amplitudes of the periodic variations. Therefore, the presence and the potential impact of the geographical and serial correlations need to be modeled and assessed for a rigorous regional solution. This is the topic to of the following section.

3. Regional Tide Gauge Model with Geographically and Temporally Correlated Disturbances and its Solution

The local trigonometric model given by Eq. (1) and (2) is extended to accommodate m local tide gauge stations data together[3] in a

[2] *The high R^2(adj) value (98.5) for the Haulover, Miami Beach, Virginia local combination is an artifact due to the large vertical reference level offset of the Virginia tide gauge data (inset in Fig. 5) dominating the total variance in the data, which is effectively removed by the reference level shift parameter in the local solution, causing an overly improved fit.*

[3] *This formulation can also be used for stitching interrupted time series, such as formulation due to the relocation of tide gauge stations, to obtain longer series (Iz and Shum 2000) or for combining tide gauge*

single solution as follows:

$$y_t^i = a_0 + \Delta a_t^i + b(t - \bar{t}) + \sum_{h=1}^{3} \left[\alpha_h \cos\left(\frac{2\pi}{P_h}\right)(t - \bar{t}) + \gamma_h \sin\left(\frac{2\pi}{P_h}\right)(t - \bar{t}) \right] + \varepsilon_t. \quad (3)$$

In this expression all stations' data are subject to the same periodic annual, semi-annual and node variations as discussed in the previous section, and a common regional trend, b. The intercept, a_0, now refers to a preselected reference station (common to all stations) from which the reference levels of the other stations are offset by an unknown amount Δa_t^i ($i = 1, \ldots, m$) also to be estimated (except one of the Δa_t^i is zero for the reference station). Hence, $m - 1$ additional unknown parameters are introduced into the formulation for the reference level differences between stations.

In most cases, as stated earlier, the discrepancies between the local trend estimates (potential biases) are not due to the data quality or the origin of the periodic variations that are different, but caused by the unmodeled systematic or transient effects in the series especially in in shorter series, which will be reduced considerably by the presence of longer series in the regional solution (Iz, 2006).

The assimilation of large number of tide gauge measurements under a single formulation also impacts the mean square error (MSE) of the solution due to the increased degrees of freedom, improving the uncertainties of the estimated model parameters and the projections that depend on these values.

A variant of the *statistical model* given by Eq. (2) takes into consideration the impact of the serially correlated (autocorrelated) disturbances for each station. The geographical correlations among the participating stations (spatial correlations) can be formulated as follows:

If the tide gauge data disturbances e_t^i at a given station i are interdependent, they can be described by *a first order autoregressive process* (known also as *red noise*) as,

$$e_t^i = \rho_i^\tau e_{t-1}^i + v_t^i, \quad (4)$$

where the correlation between e_t^i and e_{t-1}^i is $\sigma_i^2 \rho_i^\tau$ (Iz and Chen 1999). Hence, the correlation between successive tide gauge measurements' disturbances denoted by ρ_i^τ depends on time differences, denoted by τ, between the two successive epochs t and t-1, and decreases with increasing $|\tau|$.

The stochastic process $v_t^i$with the following assumed properties,

$$\mathrm{E}(v_t^i) = 0, \quad \mathrm{E}(v_t^{i^2}) =: \sigma_v^{i^2}, \ \mathrm{E}(v_t^i v_{t'}^i) = 0, \ \ t \neq t' \quad (5)$$

results in the following expressions,

$$\mathrm{E}[e_t^i] = 0, \ \ \mathrm{Var}[e_t^i] = \sigma_v^{i\,2}(1 - \rho_i^2)^{-1} =: \sigma_i^2. \quad (6)$$

Hence, using the previous three equations, the variance-covariance matrix for monthly disturbances for the i th station can be written as;

$$\Sigma_i = \sigma_i^2 - \begin{bmatrix} 1 & \rho_i & \rho_i^2 & \cdots & \rho_i^{T^i-1} \\ \rho_i & 1 & & \cdots & \rho_i^{T^i-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho_i^{T^i-1} & \rho_i^{T^i-2} & \rho_i^{T^i-3} & \cdots & 1 \end{bmatrix}, \quad (7)$$

where $t^i = 1, \ldots, T^i$, with T^i as the number of records for the i^{th} station. Note that, in the case of missing observations, the above variance-covariance matrix is generated by the time differences, τ, between subsequent tide gauge measurements.

In a regional solution with m stations, the corresponding variance-covariance matrix of the tide gauge measurements, denoted by Σ, which takes into account both within station serial correlations (autocorrelations) and among station geographical correlations is therefore;

$$\Sigma = \begin{bmatrix} \Sigma_{1,1} & \Sigma_{1,2} & \Sigma_{1,3} & \cdots & \Sigma_{1,m} \\ \Sigma_{1,2} & \Sigma_{2,2} & & \cdots & \Sigma_{2,m} \\ \vdots & \vdots & \vdots & & \vdots \\ \Sigma_{m,1} & \Sigma_{m,2} & \Sigma_{m,3} & \cdots & \Sigma_{m,m} \end{bmatrix}, \quad (8)$$

where the covariances among the stations, denoted by $\Sigma_{i,j} = \mathrm{Cov}[e_t^{i,j}]$, $i \neq j$, arise as a result of the geographical correlations among the stations and the matrices on the diagonal, $\Sigma_{i,i} = \mathrm{Var}[e_t^i]$ are for the individual station autocorrelations.

Note that all the variances and covariances reflect a stochastic relationship among station disturbances and they are not known *a priori* until the residuals from a series of local solutions are calculated and analyzed. Once this information is inferred from the local solution disturbances, the full variance-covariance matrix can be used in a generalized least squares solution in estimating the regional model parameters.

In the following section, the new regional model's effectiveness as well as the impact of the two statistical models; one based on homogeneous disturbances (solved by ordinary least squares), and the alternative, which accounts for the heterogeneous disturbances as a result of geographical and serial correlations (solved by the generalized least squares), are quantified using the tide gauge station data around the Florida Panhandle.

4. Local and Regional Model Solutions using Geographically and Serially Correlated Disturbances for Sampled Stations

Because of the varying length of the tide gauge data, only stations with long records can be used to reliably estimate and assess the

stations that are very close to each other as we did in local solutions.

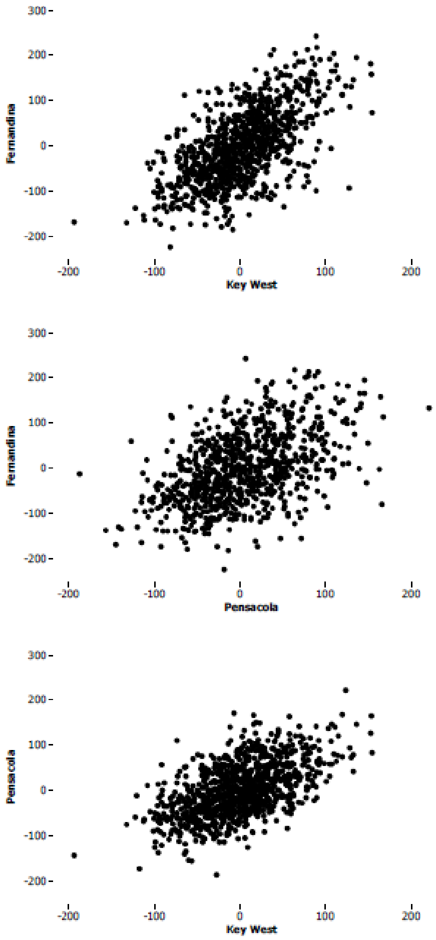

Figure 4. Geographical correlation among the Fernandina, Key West, and Pensacola stations' residuals. The correlation coefficients (on the order of 0.63, 0.48, 0.56 respectively) and the corresponding covariances listed in Tab. 3 are statistically significant (p<0.00).

serial and spatial correlations within and among stations for the variance covariance matrix of the disturbances, Eq. (4) through (8). For this purpose, ordinary least squares solutions that were carried out for the three longest stations, namely Fernandina, Key West and Pensacola were considered.

Figure 4 plots, calculated from the residuals of these solutions, clearly indicate the presence of geographic correlations among the three stations' disturbances as large as 0.6 (p<0.00) despite the large separation between the stations. On the other hand, the estimated autocorrelations (serial correlations) listed in Tab. 3 show that although the disturbances are statistically significant (p<0.00), they are not as strong as the geographical correlations among stations, because of the use of the monthly averages of the tide gauge data that tend to abase the serial dependencies.

These stations also experience sea level variations of different regions; the Atlantic Coast (Fernandina), Gulf Coast (Pensacola) and

Table 3. Variances, covariances (mm^2) and correlations (within parentheses) among three station residuals. Autoregressive residual correlation coefficient $\hat{\rho}$ for each station is also listed.

	Fernandina	Key West	Pensacola	$\hat{\rho}$
Fernandina	6430			0.3
Key West	2317(0.6)	2146		0.4
Pensacola	2171(0.5)	1522(0.6)	3342	0.5

Key West, a station in between these regions, but still exhibits geographical correlations supporting the use of common periodic components in the regional model.

The estimated serial correlations for each station and the geographical correlations among stations listed in Tab. 3 were used to construct the variance-covariance matrix of the disturbances, Eq. (8), and deployed in estimating the model parameters using the generalized least squares solution. Table 4 lists the trend estimates and their statistics for the local three stations calculated using ordinary least squares and generalized least squares solutions, the latter accounting for the serial correlations. Combined solution trend parameters (also listed in the same table) were calculated using the generalized least squares method. They are based on the amalgamated three stations' data with the model that also includes the effect of geographical correlations in addition to each station's serial correlation[4].

The local trend estimates calculated using ordinary least squares (no serial correlation) and generalized least squares (with serial correlation) turned out to be in agreement, whereas their standard errors from the generalized least squares were somewhat larger. It is well-known that omitting the serial correlations in ordinary least squares solutions has no effect on the parameters (they remain unbiased), but will decrease their standard errors (Neter et al. 1996), as also quantified in this application. All the other estimated parameters (coefficients of the periodic variations) from both solutions are in agreement within few mm.

Nonetheless, the underestimation of the uncertainties by the omission of the serial correlations are negligibly small for all three stations except for the $R^2(adj)$ values (the corresponding correlation coefficients, R, are practically the same to the first order), which may or may not have an impact in testing models. Hence, the omission of the serial correlations among station disturbances does not have an adverse impact on the estimates as well as on their statistics from this region's tide gauge data.

The consistency of the combined solutions' trend estimates obtained from ordinary and generalized least squares solutions (the latter accounts for serial as well as geographical correlations among stations) reveals that geographical correlations do not have

[4] *The regional vertical reference level in this solution is referenced to the Fernandina tide gauge station.*

VERSITA

Table 4. The trend estimates from the three longest local tide gauge series and their combined solution (regional) trend are listed. The first row of each solution are the ordinary least squares estimates followed by their standard errors (SE). The second row statistics are from the generalized least squares solution that accounts for the autocorrelation. The generalized least square solution for the combined data also includes spatial correlation among three strations. All the other parameters not listed in the table (coefficients of the periodic variations) are in agreement within few mm.

Solution	Lengthyr	Trendmm/yr	SEmm/yr	RMS Fit(mm)	R^2(adj)	R(adj)
Fernandina	112	2.03	0.07	80.4	66.4	0.81
		2.03	0.10	79.5	57.2	0.76
Key West	97	2.26	0.05	46.4	79.1	0.89
		2.26	0.08	46.5	67.3	0.82
Pensacola	87	2.12	0.07	58.0	70.3	0.84
		2.12	0.12	57.9	54.3	0.74
Combined	**112**	**2.12**	**0.04**	**67.0**	**72.8**	0.85
		2.13	0.06	67.8	59.6	0.77

an impact on the combination solution estimates and their statistics as well.

5. Final solution for regional trend using ordinary least squares

The solution results of the previous section using the three longest stations revealed that the impact of the geographical and serial correlations on the estimated parameters is negligible for the local and regional solutions for these particular data sets. Although the geographical correlations would be larger for the remaining stations because of their closer proximity to each other, their contributions to the regional solution will be considerably less than the longest series used in the previous investigation because of their shorter time span, as evidenced by their local ordinary least squares solutions. Therefore, parsimony favors the use of ordinary least squares for the final regional solution using all the available stations tide gauge data. Note that the negligible effect of omitting the serial and spatial correlation on the regional trend estimate should not be generalized to other data sets in other regions around the world. The presence of spatial and serial correlations at other regions need to be assessed using the approach carried out in this study and statistical model being proposed should be deployed if the impact of these correlations are found to be statistically significant.

The trend estimate based on the proposed combination model Eq. (3) using the ordinary least squares that make use of all 26 stations' tide gauge data (Fig. 5) is 2.14 $\pm$0.03 mm/yr. The models parameters are well separated (Variance Inflation Factors, VIF, of all the estimates are close to 1 indicative of the independency among the parameters, as listed in Tab. 5. The p–values of the estimated parameters show that all the estimates are statistically significant, except the sine coefficient of the node tide (p = 0.23) indicating that the node tide's amplitude dominated by the coefficient of the cosine term with a zero degree phase angle. The trend estimate of the regional solution and its uncertainty is practically the same as the one calculated using only three longest stations using ordinary least squares with the exception of their coefficient of determinations. The significantly smaller $R^2(adj)$ value of the all station regional solution evidence that the aggregated data is indeed regionally representative of the sea level variations around Florida Panhandle because any discordance among various stations data would have resulted in increased $R^2(adj)$ values.

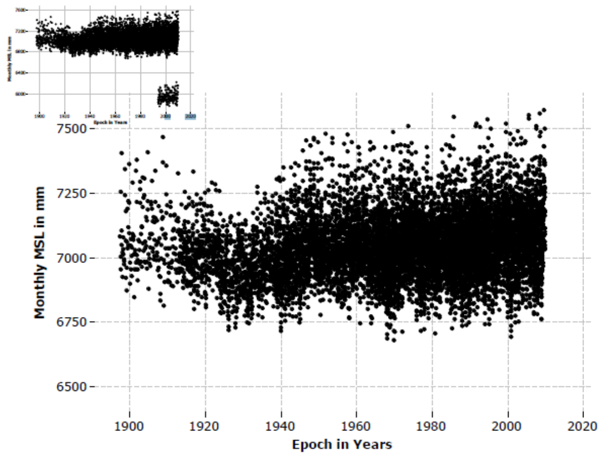

Figure 5. Combined monthly MSL data from 26 tide gauge stations. The revised local reference data still exhibit reference level differences despite the preliminary adjustments by the PSMSL. However the differences are small except one; Virginia tide gauge stations whose data are displayed in the inset.

The regional vertical reference level in this solution, realized at 1975.5, is referenced to the Fernandina tide gauge station data by virtue of being the longest series. Although this choice does not affect the estimate for the trend, the estimates for the reference level offsets vary by the choice of the reference station. In this case, the reference level offsets of all the stations are all negative with respect to the Fernandina station indicating that this station's vertical reference level is approximately 18 cm (excluding Virginia Key offset) above the other stations' reference levels. Note that all the offsets are statistically significant as indicated by their p values (p<0.00). Hence, the revised local reference data is only approximately referenced to a common reference level despite the post-processing of the metric data by PSMSL. Note that the regional

Table 5. Solution statistics for the regional model using ordinary least squares solution. Series time span: 112 years, RMS fit: 67.0 mm, and $R^2(adj)$: 87.7. The trend estimate and its standard error (SE) is in mm/yr. The other estimates and their standard errors are in mm. VIF is the variance inflation factor.

Parameter	Estimate	SE	t-value	p-value	VIF
a (intercept)	7199.12	2.08	3469.40	0.00	
b (trend)	**2.14**	**0.03**	**65.85**	**0.00**	**1.63**
α (sine node)	-1.12	0.94	-1.19	0.23	1.04
γ (cosine node)	3.88	0.97	4.01	0.00	1.05
α (sine annual)	86.36	0.93	92.91	0.00	1.00
γ (cosine annual)	9.79	0.93	10.53	0.00	1.00
α (sine semi-annual)	-40.21	0.93	-43.19	0.00	1.00
γ (cosine semi-annual)	-21.40	0.93	-23.05	0.00	1.00
Δa Apalachicola	-320.32	3.81	-84.11	0.00	1.47
Cedar Key I	-104.48	6.10	-17.13	0.00	1.15
Cedar Key II	-206.74	3.12	-66.18	0.00	1.62
Clearwater	-203.64	4.04	-50.35	0.00	1.50
Daytona Beach	-170.58	4.47	-38.16	0.00	1.24
Daytona Shores	-120.03	5.61	-21.39	0.00	1.14
Fort Myers	-233.46	3.74	-62.43	0.00	1.48
Haulover	-168.95	8.65	-19.54	0.00	1.08
Jacksonville	-102.50	5.38	-19.07	0.00	1.14
Key Colony	-221.68	5.19	-42.74	0.00	1.20
KeyWest	-60.29	2.80	-21.55	0.00	1.79
LakeWorth Pier	-234.26	11.57	-20.25	0.00	1.04
Marathon Shores	-243.11	7.10	-34.26	0.00	1.09
Mayport	-116.69	3.03	-38.58	0.00	1.62
Miami Beach	-142.13	3.45	-41.16	0.00	1.41
Naples	-196.20	3.69	-53.18	0.00	1.49
Palm Beach	-208.98	14.81	-14.11	0.00	1.02
Pensacola	-160.74	2.90	-55.51	0.00	1.73
Shell Point	-163.85	14.17	-11.56	0.00	1.03
St Marks	-188.39	7.12	-26.47	0.00	1.09
St Petersburg	-116.78	3.22	-36.24	0.00	1.62
Turkey Point	-220.28	7.34	-30.00	0.00	1.11
Vaca Key	-173.45	4.87	-35.61	0.00	1.30
Δa Virginia Key	-1304.18	5.48	-238.05	0.00	1.24

trend estimate and the amplitudes of modeled periods would be biased if the reference level offsets are not modeled in combination solutions.

The improvement in the $R^2(adj)$ value (87.7 percent) in the regional model verifies that tide gauge variations from different locations do conform with each other and support the assumption that all stations undergo the same periodic effects (Gulf and Atlantic Coast). Otherwise, any inconsistency would have decreased the $R^2(adj)$ value rather than increase. The improved coefficient of determination also reveals that local subsidence rates among stations are consistent/negligible. Again, any discordance among station subsidence rates would have had an adverse impact on the standard error the final regional trend estimate as well as on the $R^2(adj)$ value.

The plot of the residuals versus fitted values in Fig. 6 does not exhibit any dominant systematic pattern caused by the fusion of various data and suggests together with the histogram, a normal like distribution[5]. The two distinct clusters in this plot are due to the large offset in the Virginia Key station tide gauge data. The departures at the beginning and end of the straight line as revealed in the normal probability diagram in Fig. 6 are caused mostly by the 491 extreme data values out of 10380 tide gauge monthly averages.

The impact of the regional solution trend compared to the local trends as a function of t-values (Fig 3 and Tab. 6) are dramatically reflected on the projections (assuming the past and present sea level rise hold in the future) based on the new model estimates displayed in Fig. 7 and Error! Reference source not found.

The projections suggest a sea level rise of approximately 18 cm with respect to the 2011 MSL height, with a prediction uncertainty of about ±28 cm (prediction interval, PI) at 95% confidence level by 2100. Note that the prediction interval accounts also for the ran-

[5] *Nonetheless, the plot of residuals versus order (labels of observation epochs) suggest that there may still be unexplained small magnitude transient variations in the data.*

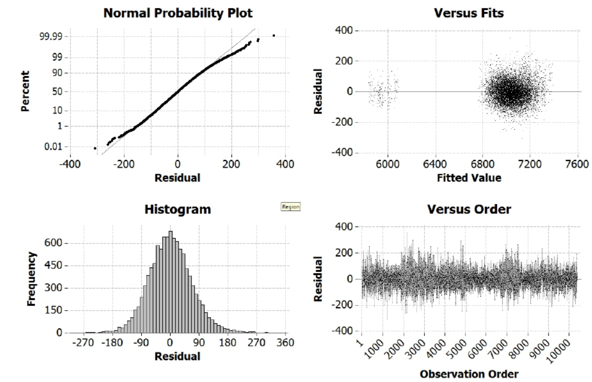

Figure 6. Properties of the residuals (mm). SE = 66.99 mm. Different depiction of the residuals all suggests homogeneous(homoscedastic) and nearly Gaussian random distribution.

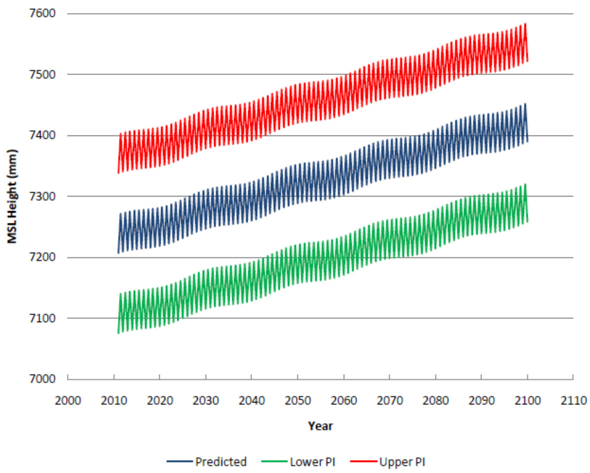

Figure 7. The predicted MSL heights (mm), plotted at 0.5 yr intervals, are referenced to the Fernandina vertical tide gauge reference level realized at 1975.5. The prediction interval (PI) is 95% confidence level. The estimated node tides are modulated by the annual and semi-annual periodicities and statistically significant. They are superimposed on the regional predicted trend.

dom variations, i.e. the residuals, and is considerably wider than the confidence interval, CI (Tab. 6). The prediction intervals shown in Fig. 7 are at least 10 fold narrower than any other projections for the region by virtue of combining all the local station data under a single regional model.

6. Validation

Currently there are only a few regional trend estimates for Florida based on the local tide gauge data to compare and validate the regional trend estimate and its standard error of this study. Most of the recent and earlier solutions use fewer number of stations because of the reliability of local trend estimates from short records.

One of these solutions as offered to represent the Florida region MSL trend, is the analysis of a single station, Key West tide gauge data, by Obeysekera et al. (2011). The study analyzed two consecutive 47 years split periods that resulted in 2.9 and 2.7 mm/yr trends (no standard errors reported). A number of local solutions by Walton (2007), and Harrington and Walton (2008) produced a range of values 1.5 – 2.4 mm/yr for the regional trend using 64 year truncated local tide gauge series (no standard errors reported). An average using the longest series in Florida calculated by Maul and Douglas as early as 1993 is 2.2 mm/yr is markedly close to the current regional estimate of 2.14 mm/yr. Yet the estimate has a large standard error of 0.5 mm/yr compared to the 0.03 mm/yr standard error of the trend estimate of this study. A multiple station study by Douglas (2001) reported 2.4 mm/yr mean trend from four stations in this region with a 0.3 mm/yr standard error of the mean.

In general, the validation of the regional trend estimate and its standard deviation requires solutions deduced from independent observations. These solutions are also to be as good as or better than the current solution so that they can be used as a baseline for comparison. Recent and earlier solutions reported above do not meet this requirement. Moreover, because all the local solutions use some tide gauge data that are also used in this study, they do not provide independent information for validation (although the results are informative to assess the use of additional data and model performance).

Satellite altimetry solutions can be considered for validation as an alternative. Such solutions are regional and fulfill the initial requirement of using data independent of local tide gauges (strictly speaking, even satellite altimetry solutions are not independent from local tide gauge data because they use local tide gauge series for calibration of the measurements).

A most recent study (Palanisamy et al. 2012) in the neighboring Caribbean area reports a regional trend estimate of 1.7 ± 0.6 mm/yr averaged over the region using satellite altimetry data during the period 1993 – 2009. This result is not statistically different than this study's estimate of 2.14 ± 0.03 mm/yr. Yet, the agreement hardly validates the current solution because the confidence interval of the regional trend inferred from satellite altimetry is markedly large.

As another alternative, a regional trend estimate is obtained using multiple satellite altimeters around the Florida Panhandle during the last 20 years as part of this study. The weighted mean of the grid trends shown in Fig. 8, where the weights are chosen as the inverse of the variance of the grid trend error estimates, is 1.5 ± 0.3 mm/yr despite the wild variations at different locations. In this case, the satellite altimetry solution is better than the neighboring Caribbean solution. However, the agreement with the current estimate using tide gauge data depends on the significance level. Moreover, the weighted average trend rate deduced from satellite altimetry data cannot invalidate the current trend estimate simply

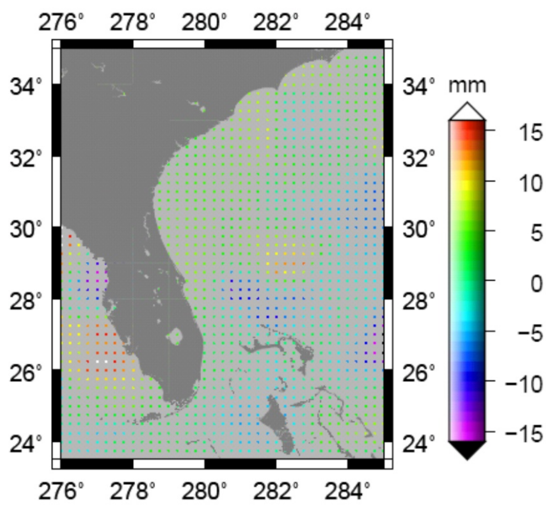

Figure 8. MSL trends (mm/yr) using multiple satellite altimetry data (∼20 yr).

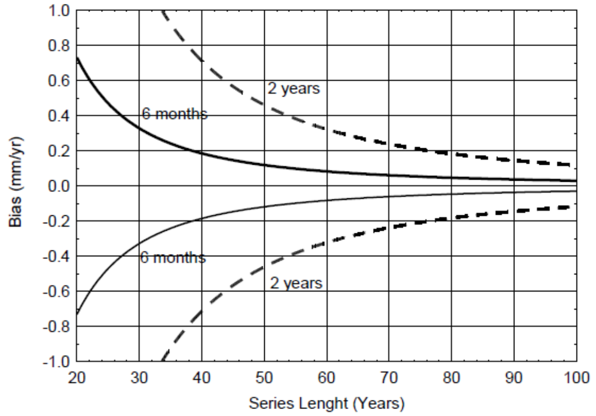

Figure 9. The trend bias induced by an unmodeled 100 mm transient sea level change that occurs at the beginning and at the end of each series of different length and lasts for 6 months and 2 years respectively.

because all satellite altimetry solutions are likely to be biased in trend estimates and in their standard errors.

The potential biases in trend estimates were demonstrated empirically in this study (see the differing local estimates displayed in Fig. 2) and theoretically, by Iz (2006). The importance of the length of the series is also emphasized by various authors (e.g. Douglas 1991) for tide gauge series. Satellite altimetry data are not an exception. Figure 9 by Iz (2006) shows that an unmodeled transient effect that lasts for two years (such as an above average meteorological condition) can bias shorter series in which such transient variations cannot be easily detected and modeled. There is no guarantee that neither the tide gauge nor the satellite altimetry data in this region are devoid of such unmodeled effects. Yet this influence decreases as the length of the series increases, hence the trend estimates from longer tide gauge data are not adversely influenced by the unmodeled sea level variations.

Furthermore, Tab. 7 (Iz 2006) lists various effects such as data quality, unmodeled or unknown transient, or systematic periodic effects acting on a 20 yr long series. They can induce trend biases as large as 0.8 mm/yr RSSE[6] (Root Sum Square Error). Such biases in trend estimate increases its variance estimate by an amount equal to the square of the bias. Again, the impact of this error on a 100 yr long series is negligibly smaller.

Yet, this is not the only problem. Current satellite altimetry solutions (including the one produced in this study) ignore the spatial and temporal correlations. The impact of the omission is different than the one on tide gauge solution examined in this study because of the close proximity of the data and the way they are processed. Satellite altimetry solutions produce gridded sea level rates interpolated from satellite altimetry data, which are then used to calculate the regional trend. Because of the omission of the spatial and temporal correlations within and among neighboring grids, the standard error of the estimated regional trend is underestimated (correlations do not impact the trend estimates). The biasing of the standard errors is usually accounted for by rescaling the error estimate upward (refer to the section on satellite altimetry in Douglas 2001). Note that between grid correlations are significantly larger than the between station correlations estimated in this study. Consequently, the error estimates are to be increased up to several folds to account for the omission of spatial and temporal grid correlations. Consequently, such rescaling makes the use of satellite altimetry solutions questionable in validating the standard error of the regional trend estimate of this study like the satellite altimetry estimate calculated in the Caribbean study.

Alternatively, a self-validation approach can be carried out using the method of bootstrapping. However, as discussed earlier, notwithstanding the validity of bootstrapping in other investigations as a statistical technique for solution and error assessment, the limitation imposed by the length the tide gauge series in this study makes such validation not so informative. If for validation, some stations from the regional solution are excluded as independent data for bootstrapping, then the regional trend estimates may or may not perform well depending on the length of the series and local sea level variations. If the local estimates are used to validate the regional solutions, then the local sea level variations induce a biased, local, and tailor made solution, with biased local trends that fit the data better than using the regional estimate, if the series are shorter. If the excluded series are long, then the regional solution does agree well with the trend of the excluded series simply be-

[6] *The effect of the node tide bias is excluded since it is estimated in this study*

Table 6. Projection statistics for the MSL heights referenced to Fernandina reference level realized at 1975.5. All units are in mm.

Epoch (yr)	Predicted Height	SE	Confidence Interval	Prediction Interval
2011.0	7191.38	2.64	(7186.20, 7196.57)	(7059.98, 7322.79)
2100.0	7373.55	4.34	(7365.05, 7382.06)	(7241.97, 7505.13)

Table 7. Contribution of the unmodeled effects to the trend estimate (mm/yr).

Error Source	Series Length (yr)		
	20	50	100
Data noise based on 3^2 and 8^2 cm^2 averaged monthly data variance.	0.3 – 0.9	0.1 – 0.2	<0.1
Periodic annual variations with ± 10 cm amplitude.	0.1 – 0.3	<0.1	-
Transient change with ± 10 cm magnitude in the MSL that last 6 and 24 months	>0.7	0.1 – 0.5	<0.1
Nodal with period 18.6 and 18 mm amplitude.	>0.8	0.1 – 0.2	-
RSSE*	>1.1	0.2 – 0.6	<0.1

* The RSSE (Root Sum Square Error) $RSSE = \sqrt{\sum_{i=1}^{4} \sigma_i^2}$ where i refers to the error source.

cause all of the three long series in our study agree well with each other and dominate the results.

A variation of the above approach is to exclude a portion of the tide gauge data that span a period of time, irrespective of which stations they belong to, from the regional solution for validation as independent data. In this case, the length of the series and their epochs become an issue. If the excluded data is at the end of the series (or at the beginning) they influence the solution as a function of the length of the excluded data as shown by Iz (2006). Otherwise, the regional solution well predicts the excluded data. Hence, this approach turns out not to be a reliable validation method because it is not randomized.

In sum, the considerable improvement in the standard error of the trend estimate using multiple station tide gauge data as in this study comes with unique challenges for its validation.

7. Conclusion

The proposed formulation is simple and evidently effective in reducing the uncertainty in the parameters of the regional sea level variability. The outcome of the proposed combination model solution is an order of magnitude improvement in the standard error of the trend estimate for the Florida region (although not much different than the combination solution using the longest three station solution). The reduced uncertainty in the trend estimates will impact testing a number of pertinent hypotheses that rely on the mean sea level (MSL) trends, including testing mean sea level acceleration in a region. Improved estimates will also lead to more accurate projections.

In the case of the Florida regional model solution, the improvements in the regional trend estimate shows that all the 26 tidal gauge data conform with each other under a single model despite the discrepancies in their local trend estimates that are caused by the short series lengths and not due to the quality of the tide gauge data. The results highlights once again, the pitfalls of using shorter series in making inferences about sea level rise as discussed in Iz (2006).

The proposed formulation is also frugal in the sense that no tide gauge data are wasted. Although the similar results can be obtained by the use of only longest series, the fusion of all tide gauge data unifies all the local vertical reference levels as defined by the local tide gauges and provides accurate statistics for the local reference level shift parameters with respect to a common reference level.

There were no significant differences between the estimated parameters and their statistics obtained from the ordinary least squares and generalized least squares solutions as a results of serial and geographical correlations, yet this result is valid only for this particular regional solution. Care should be given in modeling the correlations of the disturbances of the regional solutions among stations at localities where long series with close proximity are deployed. The proposed formulation and the two step procedure can account for their effects. In the case of using successive records with time interval less than a month, modeling serial correlations may be needed. The effect of autocorrelations can be effectively reduced by modeling local effects such as the inverted barometer effect (Iz and Shum, 2000), which are usually one of the root cause of the serial correlations.

Serial correlations may also turn out to be influential at other geographic regions such as the ones bordering the Pacific where interannual or interdecadal variability is considerably significant. The extended statistical model of the formulation proposed in this study that takes into account such variations through generalized least squares solutions is well suited in modeling the sea level variations and sea level rise in these regions.

The vigorous solution for the regional MSL variations, in this particular case for Florida Panhandle, gives the most representative solution for the region because of the contribution from various tide gauge stations, including Gulf of Mexico and Atlantic stations.

8. Acknowledgment

This study is supported partly by the DOR Research Priority Grant Program at Florida Atlantic University. We thank the two reviewers and the editors of the Journal of Geodetic Science for their time and helpful comments.

References

Barnett T.P., 1984, The estimation of global sea level change: A problem of uniqueness, J Geophys. Res., 89, C5, 7980–7988.

Cabanes C., Cazenave A. and Le Provost C., 2001, Sea level rise during past 40 years determined from satellite and in situ observations, Science, 294, 5543, 840-842 DOI: 10.1126/science.1063556.

Church J.A., Woodworth P.L., Thorkild A. and Wilson W.S., 2011, Understanding sea-level rise and variability, Wiley-Blackwell. Kindle Edition, Kindle Locations 2947-2948.

Cressie N.A.C., 1993, Statistics for Spatial Data, New York: Wiley.

Douglas B.C., 1991, Global sea level rise. J Geophys. Res. 96, C4, 6981-6992.

Douglas B.C., 2001, Sea Level change in the era of recording tide gauge. In B.C. Douglas M.S Kearney and S.P. Leatherman Sea Level Rise; History and Consequences, Chapter 3, Academic Press.

Gornitz V., Lebedeff S. and Hansen J., 1982, Global sea level trend in the past century, Science, 215, 1611-1614.

Harrington J. and Walton T., 2008, Climate change in coastal areas in Florida: sea level rise estimation and economic analysis to year 2080, Florida State Report funded by a grant from the National Commission on Energy Policy, 58 pgs.

Iz H.B., 2006, How do unmodeled systematic MSL variations affect long term sea level trend estimates from Tide gauge data journal of geodesy, 80, 1, 40-46.

Iz H.B. and Ng H.M., 2005, Are the global tide gauge data stationary in variance Marine Geodesy, 3, 209-217.

Iz H.B. and Shum C.K., 2000, Mean sea level variations in the south China sea from four decades of tidal records in Hong Kong, Marine Geodesy, 23, 4, 221-233

Iz H.B. and Chen Y., 1999, VLBI rates with first order autoregressive disturbances, J Geodyn., 28, 2-3, 131-145.

Maul G.A. and Douglas M., 1993, Sea level rise at Key West, Florida, 1846-1992: America's longest instrument record, Geophys. Res. Lett., 20, 18, 1955-1958.

Mazzotti S., Jones C. and Thomson R.E., 2008, Relative and absolute sea level rise in western Canada and northwestern United States from a combined tide gauge-GPS analysis, J. Geophys. Res., 113, C11019, DOI:10.1029/2008JC004835.

Mitchum G.T., 2011, Sea level changes in the southeastern United States past, present, and future, project report of the Southeast Climate Consortium and the Florida Climate Institute, 20pp.

Neter J.M., Kutner H., Nachtsheim C.J. and Wasserman M., 1996, Appl. Linear Stat. Mod., Richard D. Irwin, 1408.

Obeysekera J, Irizarry M., Park J., Barnes J. and Dessalegne T., 2011, Climate change and its implications for water resources management in south Florida Stoch. Environ. Res. Risk Assess., 25,. 495–516.

Palanisamy H, Becker M., Meyssignac B., Henry O. and Cazenave, A., 2012, Regional sea level change and variability in the Caribbean sea since 1950, J. Geod. Sci., 2(2), 125-133, DOI: 10.2478/v10156-011-0029-4.

Peltier W, 2004, Global glacial isostasy and the surface of the ice-age Earth: The ICE-5G (VM2) model and GRACE, Ann. Rev. Earth Planet. Sci. Lett., 32, 111–149.

Permanent Service for MSL, 2011, Home page, http://www.psmsl.org/. (Accessed on April 2011).

Tsimplis M.N. and Woodworth P.L., 1994, The global distribution of the seasonal sea level cycle calculated from coastal tide gauge data. J. Geophys. Res., 99, C8, 16,031–16,039.

Walton T.L., 2007, Projected sea level rise in Florida Ocean Engineering, 34, 2007,d 1832–1840.

Wang H. and Wu P., 2006, Effects of lateral variations in lithospheric thickness and mantle viscosity on glacially induced surface motion on a spherical, self-gravitating Maxwell Earth, Earth & Planet. Sci. Lett., 244, 3–4, 576–589.